How Computational Models Enable Mechanistic Insights into Virus Infection

Sbalzarini, Ivo F ; Greber, Urs F

Abstract: An implicit aim in cellular infection biology is to understand the mechanisms how viruses, microbes, eukaryotic parasites, and fungi usurp the functions of host cells and cause disease. Mechanistic insight is a deep understanding of the biophysical and biochemical processes that give rise to an observable phenomenon. It is typically subject to falsification, that is, it is accessible to experimentation and empirical data acquisition. This is different from logic and mathematics, which are not empirical, but built on systems of inherently consistent axioms. Here, we argue that modeling and computer simulation, combined with mechanistic insights, yields unprecedented deep understanding of phenomena in biology and especially in virus infections by providing a way of showing sufficiency of a hypothetical mechanism. This ideally complements the necessity statements accessible to empirical falsification by additional positive evidence. We discuss how computational implementations of mathematical models can assist and enhance the quantitative measurements of infection dynamics of enveloped and non-enveloped viruses and thereby help generating causal insights into virus infection biology.

DOI: https://doi.org/10.1007/978-1-4939-8678-1_30

How Computational Models Enable Mechanistic Insights into Virus Infection

Ivo F. Sbalzarini[1,2] & Urs F. Greber[3]

Affiliations:

[1]Faculty of Computer Science, TU Dresden, Dresden, Germany, [2]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany, [3]Department of Molecular Life Sciences, University of Zürich, Zürich, Switzerland

Running title: Computational Modeling of Virus Infection

Key words: modeling, simulation, computing, in-silico reconstitution, parameter fitting virus infection mechanisms, influenza virus, enveloped virus, nonenveloped virus, adenovirus, cell biology

Abstract

An implicit aim in cellular infection biology is to understand the mechanisms how viruses, microbes, eukaryotic parasites, and fungi usurp the functions of host cells and cause disease. Mechanistic insight is a deep understanding of the biophysical and biochemical processes that give rise to an observable phenomenon. It is typically subject to falsification, that is, it is accessible to experimentation and empirical data acquisition. This is different from logic and mathematics, which are not empirical, but built on systems of inherently consistent axioms. Here, we argue that modeling and computer simulation, combined with mechanistic insights, yields unprecedented deep understanding of phenomena in biology, and especially in virus infections by providing a way of showing sufficiency of a hypothetical mechanism. This ideally complements the necessity statements accessible to empirical falsification by additional positive evidence. We discuss how computational implementations of mathematical models can assist and enhance the quantitative measurements of infection dynamics of enveloped and non-enveloped viruses, and thereby help generating causal insights into virus infection biology.

Introduction

Viruses are known to infect all forms of life. They are the most ubiquitous entities on earth, exceeding 10^{30} particles, most of them bacteriophages. Viruses are multi-facetted entities at the nanoscale or microscale. They have a dual nature, the virus particle – virion, and the virus – the infected cell. Virions come in many shapes and sizes ranging from regular icosahedral particles to membrane-enwrapped amorphous entities. All viruses carry a ribonucleic acid (RNA) or deoxy-ribonucleic acid (DNA) genome, and encode their own replicases, which normally lack proof-reading activity, in contrast to cellular DNA or RNA polymerases. Error-prone replication together with genetic recombination and genomic re-assortment give rise to clouds of genetically related, but not identical viral genomes that, when packaged into particles, give rise to a so-called quasi-species of virus genomes [1-3]. This illustrates the notion that viruses act as an ensemble drawn from a cloud of related genome sequences. This is the fundamental basis for virus evolution under selection pressure, when viruses are exposed to a changing environment, for example when they infect a new host organism or cell type, or when they are under the pressure of chemicals or the immune system.

Virions unconditionally require the assistance of a cell to produce their progeny. They are obligatory parasites. Virions have to enter into a cell in order to replicate and cause an infection. This requires that pre-existing cellular mechanisms assist virus infection. These mechanisms can be explored by studying viruses, using viruses also as a proxy for understanding host-cell biology. Viruses can cause disease or be cleared by the immune system. In fact, most viruses on earth are not pathogenic to humans, since the immune system protects against foreign agents, and many viral agents simply have not coevolved with humans. However, virus infection dynamics is complex, which is reflected in the long-standing and wide-spread observation that not all cells and individuals become equally infected when exposed to the same amount of virions.

To understand how viruses cause disease, quantitative measurements of infection processes were developed that make use of genetic interference, specific drugs, and the expression of dominant-negative proteins that mimic a particular host-cell function. A classical procedure to understand mechanisms in cell and infection biology has been to use bottom-up reconstitution experiments to partially rebuild or reconstitute a certain cellular function from scratch. This comprised, for example, the actin network, the microtubule spindle, centriolar assemblies, the nuclear envelope, the endoplasmic reticulum, T-cell receptor signaling, and the motility of organelles on cytoskeletal tracks

[4]. Insights from such experiments have allowed us to reconstitute cellular processes and structures by using a small set of components from cell extracts. But, even if there is an inventory available of the proteins and factors used in a given reconstitution, and even if we know the biophysical properties of all components, we still do not understand many of the cell-based processes leading to infection until we know how these components interact with each other. This has been realized early on in the fields of cell motility and cell division, where both the inventory and the biophysics of many of the underlying components have been successfully combined into bottom-up *in vitro* systems, reproducing some key aspects of cell migration and division [5,6]. What is necessary for a system-wide analysis is that the biophysical and mechanical properties of the components are integrated into a new experimental entity to gain a more comprehensive and realistic understanding of the interaction mechanisms of infection processes.

Computational modeling and simulation can provide such an experimental entity, where mechanisms and interaction processes are reconstituted *in silico* in a fully controlled way. Experiments are then conducted computationally in what is called a “simulation”. Increasingly important aspects of infection dynamics are therefore addressed by mathematical and computational modeling [for a review, see 7]. This is particularly interesting, as it allows to estimate otherwise hidden infection parameters. This can be done in the context of an immune response or a secondary bacterial infection in the respiratory tract, for example *Streptococci* and *Staphylococci* co-infections with Influenza A virus (IAV) [8,9]. IAVs include seasonal human influenza viruses, and are found to circulate in wild water birds. They are highly transmissible, and are estimated to cause several hundreds of thousands of deaths per year [10]. Notably, the Spanish influenza pandemic had caused approximately 50 million deaths, and bacterial pneumonia was a main cause for the high lethality of the 1918 IAV pandemics [11,12]. However, up to now, the co-pathogenic mechanisms resulting in the high lethality of IAV and bacterial infections have remained unknown.

We surmise that the development of mathematical modeling frameworks of bacterial and viral infections, as well as co-infections, will help integrate progressive immuno-senescence and identify host genetic factors to advance the understanding of infectious disease to an unprecedented level of depth. In this review, we highlight some of the principles of computational modeling, and elaborate on examples for how to link *in silico* experiments with infection biology in order to enhance insight into mechanisms.

Enhancing information from cell imaging of viral infections

Advances in microscopy techniques have proven indispensable to advance insights into virus infection mechanisms in all phases of the viral life cycle, including entry, replication, assembly and egress [13-16]. Fluorescence as well as luminescence imaging further provide new opportunities for bridging *in vitro* cell culture systems to *in vivo* applications, thus aiding our understanding of virus pathogenesis and early diagnosis of viral infection and development. Fluorescence virus imaging at high spatio-temporal resolution and in super-resolution helps distinguish direct from indirect effects of anti-viral interference, for example in small-compound and RNA interference screens [17-19]. Additionally, imaging is powerful for tracking sub-viral entities, such as viral genomes, with bio-orthogonal click chemistry and for visualizing individual virion particles in cells, as pioneered with adenoviruses [20], and followed up with HIV [21,22].

Advances in light microscopy have been accompanied by developments in computational image analysis. This first included high-accuracy single-particle tracking algorithms, used for example for analyzing infection dynamics of fluorescently labeled virions at the cellular scale [23]. Single-particle tracking of fluorescently labeled virions has become a standard method of analysis, typically followed by trajectory segmentation [24-26] or the calculation of motion descriptors, such as diffusion constants and Hurst exponents. This has led to the discovery of different viral motion types, which are diagnostic of the stage in the virus entry program [27,25,26,24,28]. When combined with nanometer-precision segmentations of intracellular organelles involved in virus entry, such as endosomes [29], imaging and tracking of individual virion particles enables spatial statistical studies of how virions position in a cell with respect to those structures. This can, for example, be used to derive interaction maps that explain the action of a drug or a genetic perturbation [30]. Furthermore, the development of state-of-the-art correlation techniques involving electron microscopy with nanometer-precision localization of components, and fluorescence microscopy with larger context of the infected cell has enhanced insights into virus morphogenesis with unprecedented ultra-structural detail [31].

A single high-resolution image of fluorescent virions already contains a lot of information. For example, spatial statistics, such as Ripley's K-function, can be used to decide whether the virions are uniformly distributed, or clustered [32]. This provides direct evidence about their interactions, even if mediated by confounding factors. If a second color channel shows a host-cell structure of interest, such as endocytic compartments, a generalization of spatial statistics can be used to infer the most likely

interaction between the virions and the host-cell structure [30]. This type of analysis estimates an interaction potential that is most likely responsible for the observed distribution of virions with respect to the given distribution of host-cell structures. As still images suffice for these analyses, the procedure naturally applies to single-molecule localization modalities, such as PALM and STORM [33]. And, since the gradient of the estimated potential can be interpreted as a force, the expected dynamics of the virions can be estimated for the next time step, i.e., a single still image can be used to predict how the virions are likely to move in the next time instance.

Mechanisms of infection

The development and manifestation of infectious diseases is highly complex and involves host cells and pathogens in all possible combinations. Two principle types of approaches are used to analyze host-pathogen interactions: top-down and bottom-up. Top-down approaches describe the larger context of infectious disease phenotypes, and involve epidemiology, physiology, and omics measurements, including messenger RNAs, proteins, lipids, metabolites, and sugars at the level of an individual, an organ, or a particular cell type. When combined with microscopy, these measurements increasingly enable phenotypic profiling at the subcellular level, for example by using single RNA detection assays or imaging mass cytometry [34,35]. Methodological refinements further allow quantitative assessments and data correlation.

To stringently support new concepts, mechanistic insights are, however, required. At the level of cells, a virus infection can be dissected into distinct steps, for which mechanisms can be elucidated [36,13,37]. These mechanisms can then be considered bottom-up, that is from their constituting biochemical and biophysical processes. For example, the mechanism by which a virus particle binds to cells involves one or several receptors that directly bind to the virus and hence initiate infection. In addition, binding of virions to cells may involve attachment factors, which bind to the particle, but do not lead to infection in the absence of the receptor [38]. To further complicate matters, the virion binding to cells might be tuned by facilitating proteins, which do not bind the virion, but indirectly enhance infection, for example through cell signaling and upregulation of the receptor levels on the cell surface [39].

Regardless of whether a top-down or a bottom-up approach is chosen, mechanistic insight typically starts from correlative observations. They always incorporate different layers of evidence based on a robust and quantitative observation methodology. Ideally, they are predictive and transferable to other systems. At the molecular level, they can

be interrogated by interference using knock-out approaches or more subtle changes in the suspected molecules. A classical approach for the identification of critical host factors in virus infection has been genome-wide RNA interference. Initially, screens were conducted with unmodified double-stranded synthetic interfering RNA (siRNA) complementary to a given host mRNA that was knocked down through the cellular Ago-RISC-dependent silencing complex [40]. Subsequent, screens used chemically modified passenger strands and 5' overhanging nucleotides to enhance knock-down specificity and reduce off target effects [41]. RNAi screens were conducted against a wide range of virus-infected cells, including Influenza A virus, vaccinia virus, bunyavirus, adenovirus, herpesvirus, rhinovirus, rotavirus [42-46].

Although off-target effects of RNAi have limited the interpretation of such screening data [41], the power of large-scale RNAi screens was found to be significantly enhanced by systematic analyses of a range of different pathogens and siRNAs with different chemical properties, notably in combination with a Parallel Mixed Model (PMM) approach to enhance the statistical power of hit detection using parallel screening [47]. PMM allowed the inclusion of siRNA weights that could be assigned according to available information on RNAi quality. Moreover, PMM enhanced the predictability of hits for follow-up screens through the determination of a sharedness score. This enabled the identification of novel hit genes involved in the entry pathway of most of the pathogens in the study. Recently, genetic screens were reported for cell infection with picornaviruses, a large family of positive-sense RNA viruses with severe impact on human health. In a genome-wide haploid loss-of-function screen, the phospholipase PLA2G16 was found to be an essential host factor for rhinovirus infection by supporting the translocation of the viral RNA genome from endosomes to the cytosol [48]. The same factor was picked up for rhinovirus infection in a genome-wide forward screen using a murine haplobank [49]. Arguably, although elegant, genetic loss-of-function screens for infection are limited to genes that are non-essential for host-cell survival. Computational modeling of the infection efficiency from toxic loss-of-function phenotypes could therefore enhance the breath of genetic haploscreens.

Computational modeling in virology

Logic and mathematics provide strong foundations for modeling of biological phenomena. Modeling is the intellectual process of formalizing knowledge about a system or a process. A model constitutes a hypothesis of how one believes things could work. Models can be extracted or learned from data or constructed from known

biochemical and biophysical evidence. Data-driven (also called “top-down”) models formalize patterns and correlations in the data that are extracted using methods from statistics or machine learning. Examples range from correlation analyses to reconstructing molecular interaction networks derived from high-content screening datasets [50,47] to classifying viral motion types using machine learning [24]. Data-driven models suggest mechanisms, and can be used to show necessity of a process or molecule in a perturbation experiment.

Mechanistic (or “bottom-up”) modeling aims to reconstitute a process or a system from known fundamental principles of chemistry and physics, such as conservation of mass or the statistical mechanics of chemical kinetics. Bottom-up models are akin to *in vitro* reconstitution experiments with the important difference that all system parameters can be controlled and the exact physics and chemistry assumed is known. As such, these models can be used to show sufficiency of mechanisms, which is more powerful than only showing sufficiency of ingredients, for example in an *in vitro* reconstitution [51].

Models of both kinds are then studied in simulations. A simulation is an experiment performed on a model. Computer simulations enable us to leverage the power of modern electronics in order to simulate models of unprecedented complexity and level of detail. A simulation allows any part or parameter of a model to be systematically perturbed or altered, and high-performance computers can simulate hundreds of thousands of model perturbations in a short time. Simulations also provide access to dynamic data, while experimentally end-point assays are often used. Using the end-point datasets to build or identify a model that reproduces them, and then using that model to predict the dynamics of how the system transitioned from its starting point to the observed end-point, helps interpret biological information and carve out the essential mechanisms.

Bridging the gap - from observations to mechanisms by computer simulations

Quantitative image analysis, combined with structural and biochemical data, provides a wealth of information that can be used to build models of the chemical and physical mechanisms of infection for different viruses. The model formalizes a hypothesis. It captures our current understanding based on the available information. An important question therefore is how to validate, test, and further refine the model. One approach is to test if the model is necessary and sufficient to explain a process. Showing necessity is mostly done in perturbation experiments. If the process stops working upon knocking

out or altering a molecular component that is predicted to be present according to the model, then we know that it is necessary. Sufficiency is mostly shown in reconstitution. This includes *in vitro* reconstitution or *in silico* reconstitution, that is, computer simulation. If a reconstitution of the model, which only consists of known and controlled components, reproduces the correct behavior, we know that the components are sufficient. In simulations, not only the components but also their interactions and the assumed laws of physics can be freely reconstituted. Simulations therefore bridge the gap between observation and mechanism and show sufficiency of a mechanism by *in silico* reconstitution.

Simulations at different levels of detail

A rich landscape of modeling and simulation has been developed over many years, ranging from atoms to continua. When applied to viruses, atomistic molecular-dynamics simulations have mostly considered the capsid [52]. The basis for molecular-dynamics simulations are often high-resolution cryo-EM tomograms, for example of the HIV capsid [53]. These structures provide the initial placement of the atoms in the simulation, which then gives insight into the atomistic dynamics over time, for example capsid dissolution [54].

Alternatively, simulations consider the *de novo* self-assembly of theoretical capsid structures, such as idealized polyhedral structures, and explain the thermodynamics of their assembly [55]. For some viruses, like tobacco mosaic virus (TMV), complete all-atom simulations of the entire virus, including capsid and RNA have been performed [56]. Such all-atom molecular simulations are, however, costly and limited to short time scales of nanoseconds to milliseconds. The above all-atom study of TMV for example simulated the time-resolved dynamics of 1 million atoms over 50 nanoseconds life time. Studying larger entities or longer processes, such as virus entry and virus-receptor binding, necessitates simplifications, such as coarse-grained methods where multiple atoms are lumped together. For example, this approach has been used to simulate HIV capsid shape and investigate capsid stability [57], and to study the dynamics of several viral structures, including the full satellite tobacco mosaic virus (STMV) particle, the satellite tobacco necrosis virus (STNV) capsid, poliovirus capsid, and the reovirus core [58].

Structural data derived from atomistic or coarse-grained models are often only available for isolated time points and do not represent the entire dynamic process, such as capsid assembly and maturation. The reason is that full molecular dynamics trajectories are

computationally too expensive to obtain. In such cases, elastic network models can be used to interpolate between structural states by assuming that the molecular constituents or coarse-grained particles are connected to each other by elastic springs. The so-obtained elastic structure can then be computationally morphed from one structural state to another, providing energetically plausible molecular trajectories. This approach has, for example, been used to study bacteriophage HK97 capsid maturation [59] and conformational changes in hepatitis C virus helicase [60].

Further coarse-graining models, lipid membranes can be described as continuous elastic sheets and viruses as rigid polyhedral arrangements of receptor binding sites. Such models have been used to simulate virus-receptor binding in elastic diffusive membranes. The corresponding simulation method, termed Brownian Reaction Adhesion Dynamics (BRAD), was first applied to study HIV attachment [61]. The approach was then extended to simian virus 40 (SV40) and compared with high-resolution experimental data, highlighting the importance of in-membrane receptor diffusivity for efficient attachment of SV40 to host cells [62].

While such simulations are feasible for small numbers of virions, they are computationally intractable for large virion concentrations. Once the number of virions exceeds a few thousand, individual virions cannot be represented explicitly any more. In this case, the density or concentration field of virions is modeled as a continuous distribution, leading to completely continuous descriptions. This has, for example, been used to simulate the spreading of human adenovirus across epithelial monolayers, simulating what amounts to hundreds of thousands of virions and thousands of cells [63]. Continuous models have also been used to describe the intracellular trafficking of adenoviruses in host cells using diffusion-reaction-advection equations that also account for the intrinsic dynamics of the microtubule network [64].

Taken together, simulations of viral structures, including capsids, envelopes, genomes, and surface proteins have implications for many fields of study, ranging from atomistic models to cell biology, imaging, and anti-viral therapeutics. Despite this importance, concerted community efforts of creating standardized and portable simulation software frameworks are scarce and as of now, have been limited to specific applications [65]. A generic framework of how computational modeling and wet lab experimentation complement each other is illustrated in Figure 1.

Types of computational models

Computational models in virology can be classified along five axes – 1) discrete vs. continuous, 2) spatiotemporal vs. temporal, 3) stochastic vs. deterministic, 4) hypothesis-driven vs. data-driven, and 5) white-box vs. black-box models.

Discrete vs. continuous models

Examples of discrete models include atomistic and individual-virion simulations, where atoms or virions are modeled as discrete entities. In continuous models, the individual entities are not separately represented, but only their density or distribution in space and / or time is tracked. Continuous modeling approaches have been applied to an important question in virus assembly, namely, how does a virion become infectious? In the case of HIV, where proteolytic maturation is key to gaining virion infectivity, it has been hypothesized that the cleavage of the matrix (MA) domain from the envelope (ENV) domain (comprising the viral glycoprotein) and the spreading of MA in the virion allows ENV to loosen up and cluster the trimers for assembling a functional fusion machinery in the infectious virions [66]. Indeed, reaction-diffusion models were used to decouple MA from ENV, and thereby simulated an aspect of virion maturation [67]. Since assembly and maturation are coupled events, and the transition of immature to mature capsid requires conformational changes in capsid (CA), researchers also used coarse-grained discrete models for simulating lattice molecular assembly and non-diffusional curling vs. *de-novo* assembly [68,69]. The ultimate aim here will be to describe the cleavage of the hetero-polymer as a cascade of events, and predict to what extent the cleavage has to occur in order to yield functional capsid assemblies.

Spatiotemporal vs. temporal models

Spatiotemporal models explicitly represent the spatial localization or distribution of viruses, such as done in most of the above-mentioned examples. In contrast, temporal models track the dynamic evolution of an aggregated quantity, such as the total virus load or multiplicity of infection, without reflecting its spatial localization. This is traditionally the case in viral kinetics models [70]. Multi-scale spatio-temporal modeling of virion maturation has been performed by Markov models where the free energy landscape of intermediate states was averaged [71]. One can expect that the combination of molecular dynamics and Brownian dynamics models will provide more computational cost-efficient simulation results.

Stochastic vs. deterministic models

Another distinction is whether a model is stochastic or deterministic. In stochastic models, certain events happen probabilistically, such that the evolution of the infection state cannot be accurately predicted, but probabilities of different evolutions can be evaluated. For example, the infection probability of cells depends on the local virion concentration, and this has been implemented into a 'white box' model for the simulation of infection spread in a tissue culture model [63]. Another example for stochastic modeling in virus infection has been the trafficking of individual virions along microtubules, in an attempt to better understand the number and kind of cellular motor proteins involved in periods of virion motion bursts in the cytosol [72]. Stochastic simulations of fluorescent adenovirus particle motions involved an energy function and known parameters of motor stepping and on/off rates on microtubules. They predicted that one to two motors are bound per virion during an active motion burst. The model accurately reproduced the virion motions from live-cell imaging data. It predicted that the major capsid protein, hexon, was the receptor for the dynein/dynactin motor complex [72]. This notion was co-incidentally confirmed using biochemical pull-down and infection assays [73].

In contrast to stochastic models, in a deterministic model all events happen with certainty, which typically requires complete knowledge of the molecular mechanisms at play. Deterministic models have been implemented for aspects of actin polymerization, and were extrapolated to acto-myosin-based cell motility [5,74]. Deterministic models might be implemented for simulating the disruption of a non-enveloped virion, where the interactions between capsid proteins are known in atomic detail. Such simulations would be informative to predict, for example, if mechanical forces acting on virus particles during virion drifting motions on the cell surface are sufficient for the partial disruption of the virion during entry, as observed in the case of adenovirus [75-77,26,78]. The model would implement information about protein-protein contacts from the crystal or cryo-EM structures and the anisotropic mechanics of the icosahedral particle measured by atomic force microscopy [79-82]. It would inform about the force that is needed to pull out a capsomer at the virion vertex. Such information has high relevance, since the acto-myosin filaments mediating the virion drifts on the cell surface are much larger assemblies than the virion itself, and the minimal components of this machinery are unknown. How many motors, how many filaments are involved? Modeling might provide information about the organization of the cytosolic region

proximal to the plasma membrane, for example in relation to the picket–fence model [83].

Hypothesis-driven vs. data-driven models

Hypothesis-driven models are formulated based on an expectation or suspicion for which no data need be available at first. First, a hypothesis is formulated and then formalized, for example in the form of mathematical equations, rule sets, or chemical pathways. The resulting model is then simulated in order to study its behavior in an attempt to falsify the hypothesis by comparison with observations and known facts.

Alternatively, models can be learned from data without formalizing a hypothesis. This is often useful in the initial exploratory phase of a study, or when seeking higher-order patterns in data that are not apparent to the human observer. While methods of statistical analysis have done precisely this for a long time, recent breakthroughs in machine learning and artificial intelligence have brought a new quality to data-driven modeling. Modern machine-learning methods, such as deep and convolutional neural networks, are exceedingly powerful at discovering patterns in complex datasets. As they are not hypothesis-driven, they do not directly serve the purpose of showing a biological mechanism, but they uncover correlations and open the possibility to classify previously unseen data. Supervised machine-learning requires large amounts of training data from known conditions in order to learn the correlations. Therefore, it is not surprising that some of the first applications in virology were in detecting correlations of host-cell gene expression levels with viral infection status, for example in hepatitis B virus infections [84].

Another application where training data are typically available is in single-virion tracking experiments. Using automated tracking software [23], hundreds or thousands of trajectories can be automatically extracted from large image datasets and different motion patterns can be labeled by hand. From these training data, a supervised machine-learning method can then learn the descriptive features of the motion patterns and predict them in new trajectories as well [24]. A third example of a successful application of machine-learning to virology is the discovery how cell-to-cell variability influences virus infection by endocytosis [85]. This was made possible by analyzing large sets of high-content screening images and learning models that link the cellular context in an image to the observed infection dynamics.

White box vs. black box models

Models are also classified according to the number of free parameters they have, which typically are fitted to experimental data. White-box models have no or just a few unknown parameters. A white-box model is the most direct evidence for sufficiency of a mechanism. If, for example, all diffusion constants, infection probabilities, and binding affinities are independently measured, and the model recapitulates the data, this is strong evidence that the modeled mechanisms are sufficient. In contrast to white-box models, black-box models are entirely identified by parameter fitting. They therefore provide indirect ways of estimating quantities that are not directly measurable or observable once the basic mechanism is known. They do, however, always leave some ambiguity about the actual mechanism, as different mechanisms could recapitulate the same data for different parameter values.

Enabling statistical inference

An important role of simulations is to enable statistical inference from experimental data. Inference is typically done in either the maximum-likelihood framework or the Bayesian framework. Both require a “forward model” of the observation process. Assume, for example, that one would like to infer how the morphology of endosomes changes upon virus entry. Since the question is about dynamics, live imaging is desirable. Light diffraction, however, limits the details visible in the images, especially when observing small structures like endosomes [29]. Using a simulation to model the image-formation process in the microscope, however, one can robustly infer the time-resolved endosomal shape that is most probable to have created the observed image [86]. Likewise, a simulation of virus plaque formation in infected tissues can be used to infer the mechanism of virus spread that is most likely to lead to the observed plaque dynamics [63].

Extracting information from merged heterogeneous complex data

Increasingly, viral processes are studied by combining different sources of data, such as fluorescence microscopy, electron microscopy, biochemical assays, infection assays in tissue culture, structural data, and epidemiological data. These data are heterogeneous, as they come in different forms, such as images, numbers, graphs, and time series. Merging complex heterogeneous data in order to extract information from them usually requires a computer model of the studied process. The simulation environment provides

a uniform container into which all data types can be fused, as long as they can be computationally handled.

Aiding experimental design

Simulation models almost always contain parameters, such as diffusion constants, reaction rates, or binding affinities. While some of them can be measured experimentally, it is usually undesirable to blindly measure all of them. Instead, one wishes to focus experimentation on parameters that are important to the overall behavior of the model. Once a simulation of the model is available, global sensitivity analysis methods [87] can be used to determine parameter importance. Experimental measurements or perturbations can then focus on those parameters that are predicted to be important for the function of the modeled process. At the same time, parameters that turn out to have little or no influence on the model behavior can be removed from the model, hence simplifying the model. Models thus become evolving hypotheses that suggest both next experiments and iterative refinements by incorporating the experimental results.

Fitting the values of unknown parameters by design centering

Data fitting is the standard approach to using a model to infer unknown values of, for example diffusion constants or reaction rates [88]. Almost invariably, the task of model fitting is formulated as an optimization problem. This is to find the parameter values, for which the model output is as close as possible to the experimentally measured data. However, optimal fitting can be dangerous for two reasons: First, the model necessarily is an incomplete approximation to reality and the experimental data include unknown measurement uncertainties. The best fit of one to the other is not necessarily the most meaningful in reality. Second, optimizing the fit may lead to models of growing complexity that reproduce intricate details or trace meaningless measurement errors, obscuring the basic mechanism. This is known as overfitting in machine-learning approaches.

Instead of formulating parameter inference as an optimization problem, it can be formulated as a design-centering problem. Design centering is a classic problem in engineering, first described in the electronics community [89]. In design centering, one specifies criteria that define a good model. These criteria can for example be that all measurements are matches within 1% error, and all concentrations have positive values. Any set of parameter values for which the model fulfills these criteria

corresponds to an acceptable model, of which there are usually many. Design centering now finds the one model/parameters that are acceptable *and* have maximal robustness against random fluctuations in the data or the parameters. That is, the final model has the highest probability of still being acceptable for the next, yet unseen experiment and slight changes in the parameters only minimally alter its behavior.

Since design centering finds robust models that fit 'well enough', it is free of overfitting and naturally generalizes across experimental conditions. Particularly in virology this viewpoint intuitively makes sense, since the robustness of an infection mechanism against, for example changing immune response and changes in the biophysical parameters of the cell is evolutionarily selected for. It is thus expected that design-centered mechanisms have higher chances of surviving and are thus more likely to be true. Importantly, design-centering can therefore also be used to select between different competing models and choose the more likely one, since it naturally quantifies the robustness of a model.

Despite its advantages, design-centering is only rarely used. The reason is the high computational cost it incurs. In fact, design centering has been proven to be non-deterministic polynomial (NP)-hard, which means that it is impossible to be solved efficiently on a deterministic computer [90]. Recently, however, the first efficient approximation algorithm for general design-centering problems has become available [91]. This procedure may replace optimization when fitting model parameters to data, ultimately leading to more robust models that account for measurement uncertainties and that are evolutionarily plausible.

Modeling influenza A virus (IAV)

With regards to enveloped viruses, emerging experimental data from lipidome analyses, together with cryo-EM structures have motivated computational efforts to generate glycolipid-protein interaction maps, and to explore if a particular protein of the virion envelope is directly exposed to the environment and might hence present a direct drug target [92]. Mapping of the potential impact of the glycan residues on the viral glycoproteins might inform about the shielding of particular antigenic sites, with a possible impact on antigenic variation influenced by the kind and extent of glycosylation, which is a key issue in the design of effective vaccines against influenza virus.

With respect to IAV, an interesting regulatory protein is the M2 ion channel. M2 is a homo-tetrameric protein with a single transmembrane segment each. It is present in the

virion envelope and in membranes of the secretory and endocytic pathways of infected cells. M2 is well known for its proton conductance [93-95]. It is widely conserved among IAV, indicative of important function in the viral life cycle. In fact, if the M2 channel is blocked by drugs, such as amantadine, virus entry is inhibited [96]. In absence of amantadine, the interior of the virion acidifies when located in a low-pH endosomal compartment, and the ribonucleoprotein (RNP) complexes dissociate from the capsid-coat protein M1 [97,94,98,99]. Modeling enabled elucidating the mechanism of proton conductance, involving histidine imidazole-imidazole stabilization of the charge in the lipid bilayer [100]. The tetrameric nature of the M2 channel thereby helps distribute the positive charge across different histidine residues, and thereby helps minimize a futile cycle in order to favor the productive cycle of proton conductance.

IAV rapidly evolves resistance against amantadine, and modeling again helped explain how this is possible. Using molecular dynamics simulations, it has been found that a single point-mutation (Ser31Asn) in M2 rendered IAV resistant to the M2-channel blocker amantadine [101]. Surprisingly, amantadine still bound to the S31N mutant of M2, more flexibly than in the wild-type channel, in which it stably binds to the plugging region [102]. In the mutant configuration, water surrounding the drug can easily transport protons past the plugging drug, thereby explaining proton transport even with the drug bound to M2. This sufficient mechanism has later been confirmed by a combination of nuclear magnetic resonance (NMR) experiments and simulation data using rimantadine, an anti-viral compound structurally related to amantadine [103].

By neutralizing the pH in acidic cellular compartments, M2 also subverts the normal function of endosomes and of the Golgi apparatus, and it inhibits premature conformational changes in the newly synthesized viral HA protein in the Golgi. During virion budding from the plasma membrane, M2 replaces the ESCRT (endosomal sorting complexes required for transport) machinery by localizing to curved membrane domains. In simulations reconstituting the action of the M2 channel in membrane budding, an excess of lipid moieties was necessary for obtaining a reliable representation of viral budding, due to the wedging effects of M2 and lipid bilayer curvature [100].

Besides conducting protons, M2 is thought to also conduct Na^+ and K^+ ions [104-106]. This feature is important for virion uncoating and infection, as shown in acid-bypass experiments where extracellular IAV particles attached to the plasma membrane gained infectivity when exposed to millimolar extracellular concentrations of K^+ [107]. Na^+ and

K^+ conductance through M2 is possible, despite the clear preference that M2 has for H^+ over Na^+ [108]. The reason why this is possible is that the concentrations of Na^+ or K^+ in endo-lysosomes are 5 to 6 orders of magnitude higher than the proton concentration [109]. In liposome reconstitution assays, M2 was shown to be slightly permeable to Na^+ and K^+ [102]. Thus, it is possible that H^+, Na^+ and K^+ have similar fluxes across M2 when the virion is in an acidified endosome. Since the virion is in a Na^+-rich environment when infecting cells from the outside, it will be particularly interesting to model the flux of K^+ across M2 as a function of the K^+ concentration, and to simulate possible effects of K^+ on the M1-RNP interactions or other components of the virion lumen, such as the viral RNA-dependent RNA polymerase.

Outlook

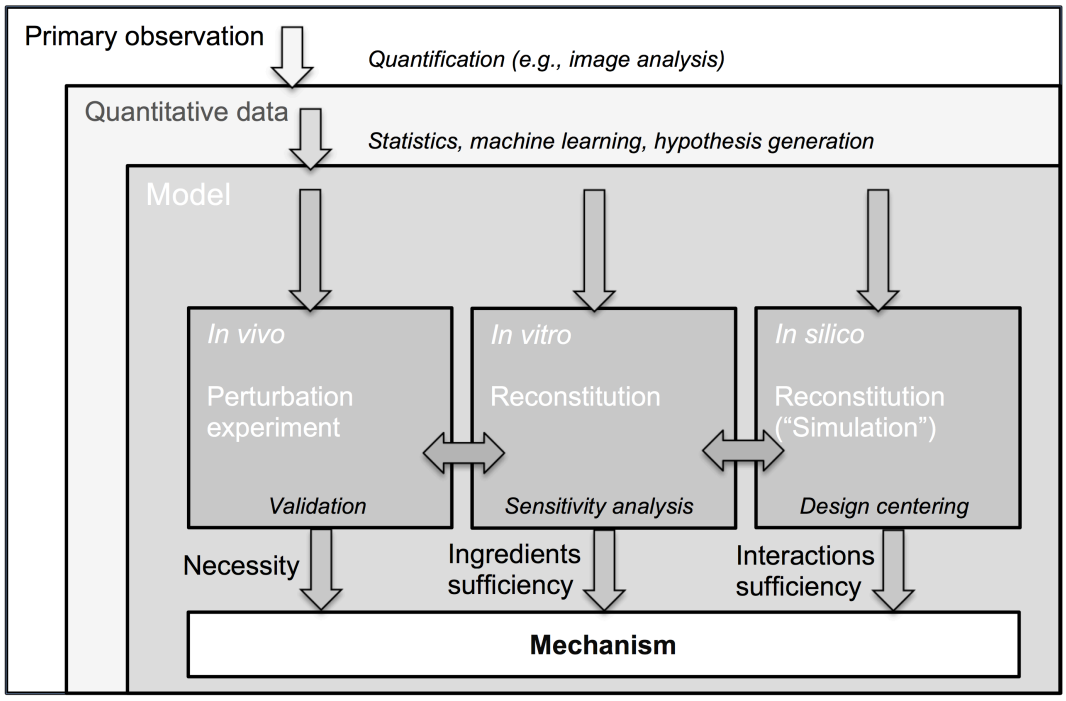
Computational methods have come a long way in virology. They range from now-standard computational image analysis, such as single-virion tracking [23], to machine-learning approaches for automated model extraction [24], and to identifying predictive interaction potentials between virions and host-cell compartments from images [30]. Many of these developments were inspired by applications in virology in the first place, rendering it a truly interdisciplinary effort. Computer simulations of learned or hypothesized models enable *in silico* reconstitution, and can show sufficiency of a mechanism, rather than just of a list of ingredients or molecules. Simulations can be done at different levels of resolution, from all-atom molecular dynamics simulations to continuum models.

On the molecular scale, it seems obvious to take advantage of the increasingly detailed structural information on inter-atomic contacts in order to model viral capsid mechanics. Discrete mechanical assembly models can be compared with continuum thin-shell descriptions in order to disentangle stochastic and deterministic mechanisms. Eventually, this may lead to the simulation of conformational changes in virus particles, as for example triggered by uncoating cues [110,38,111]. In a next step, the interaction of the exposed membrane-active viral proteins with host-cell membranes can be modeled using a coarse-gained approach. This could be rewarding, and inform about how higher-order oligomers lead to enhanced ability to disrupt membranes, for example by recruiting additional monomers from the surface without a kinetic barrier of membrane insertion for the additional monomers. Such mechanism has been proposed for the disruption of bacterial membranes by antimicrobial peptides [112]. It could have implications on the mechanism of viral membrane rupturing proteins, such as protein VI

of the human adenovirus, which preferentially binds to and disrupts ceramide-rich lipid bilayers [113,114]. We anticipate that depending on the oligomer formed, these membrane-active peptides make different sized channels, which can be measured by various sized dextrans, or thermodynamically disrupt the membrane as proposed for antimicrobial peptides. Experiments could be complemented by continuum simulations of membrane leakage. In addition, structural virion information can be used for predictive modeling of flexible regions of a virion, including intrinsically disordered domains on the surface of the particle. These regions are prone to interact with many different proteins in unpredictable manners, and often represent antigenic sites for the binding of neutralizing antibodies. Including these effects would greatly enhance the currently rather crude multi-scale models of virion attachment and in-membrane receptor diffusion [61,62].

Despite exciting prospects, we are aware of the difficulty and the challenges in modeling biological phenomena, including viruses. In this regard, we would agree with George Edward Pelham Box that ‘all models are wrong, but some are useful’, and knowing this with Richard Feynman – ‘it is much more interesting to live not knowing than to have answers which might be wrong.’

Figure 1: An integrative modeling and wet-lab approach towards mechanisms of infection biology



References

1. Eigen M (1971) Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften 58 (10):465-523

2. Domingo E, Sabo D, Taniguchi T, Weissmann C (1978) Nucleotide sequence heterogeneity of an RNA phage population. Cell 13 (4):735-744

3. Domingo E, Martinez-Salas E, Sobrino F, de la Torre JC, Portela A, Ortin J, Lopez-Galindez C, Perez-Brena P, Villanueva N, Najera R, et al. (1985) The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance--a review. Gene 40 (1):1-8

4. Liu AP, Fletcher DA (2009) Biology under construction: in vitro reconstitution of cellular function. Nat Rev Mol Cell Biol 10 (9):644-650. doi:10.1038/nrm2746

5. Pollard TD, Cooper JA (2009) Actin, a central player in cell shape and movement. Science 326 (5957):1208-1212. doi:10.1126/science.1175862

6. Pollard TD (2017) Nine unanswered questions about cytokinesis. J Cell Biol. doi:10.1083/jcb.201612068

7. Boianelli A, Nguyen VK, Ebensen T, Schulze K, Wilk E, Sharma N, Stegemann-Koniszewski S, Bruder D, Toapanta FR, Guzman CA, Meyer-Hermann M, Hernandez-Vargas EA (2015) Modeling Influenza Virus Infection: A Roadmap for Influenza Research. Viruses 7 (10):5274-5304. doi:10.3390/v7102875

8. Chertow DS, Memoli MJ (2013) Bacterial coinfection in influenza: a grand rounds review. JAMA 309 (3):275-282. doi:10.1001/jama.2012.194139

9. Leimer N, Rachmuhl C, Palheiros Marques M, Bahlmann AS, Furrer A, Eichenseher F, Seidl K, Matt U, Loessner MJ, Schuepbach RA, Zinkernagel AS (2015) Nonstable Staphylococcus aureus Small-Colony Variants Are Induced by Low pH and Sensitized to Antimicrobial Therapy by Phagolysosomal Alkalinization. J Infect Dis. doi:10.1093/infdis/jiv388

10. Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, Fukuda K (2003) Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA 289:179 –186

11. Johnson NP, Mueller J (2002) Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. Bull Hist Med 76 (1):105-115

12. Taubenberger JK, Reid AH, Fanning TG (2005) Capturing a killer flu virus. Sci Am 292 (1):48-57

13. Greber UF, Way M (2006) A superhighway to virus infection. Cell 124 (4):741-754. doi:10.1016/j.cell.2006.02.018

14. Brandenburg B, Zhuang X (2007) Virus trafficking - learning from single-virus tracking. Nat Rev Microbiol 5 (3):197-208

15. Smith GA, Enquist LW (2002) Break ins and break outs: viral interactions with the cytoskeleton of mammalian cells. Annu Rev Cell Dev Biol 18:135-161

16. Radtke K, Dohner K, Sodeik B (2006) Viral interactions with the cytoskeleton: a hitchhiker's guide to the cell. Cell Microbiol 8 (3):387-400

17. Burckhardt CJ, Greber UF (2009) Virus movements on the plasma membrane support infection and transmission between cells. PLoS Pathog 5 (11):e1000621. doi:10.1371/journal.ppat.1000621

18. Mothes W, Sherer NM, Jin J, Zhong P (2010) Virus cell-to-cell transmission. J Virol 84 (17):8360-8368. doi:JVI.00443-10 [pii]; 10.1128/JVI.00443-10

19. Florian PE, Rouille Y, Ruta S, Nichita N, Roseanu A (2016) Recent advances in human viruses imaging studies. J Basic Microbiol. doi:10.1002/jobm.201500575

20. Wang IH, Suomalainen M, Andriasyan V, Kilcher S, Mercer J, Neef A, Luedtke NW, Greber UF (2013) Tracking viral genomes in host cells at single-molecule resolution. Cell Host Microbe 14 (4):468-480. doi:10.1016/j.chom.2013.09.004

21. Sakin V, Hanne J, Dunder J, Anders-Osswein M, Laketa V, Nikic I, Krausslich HG, Lemke EA, Muller B (2017) A Versatile Tool for Live-Cell Imaging and Super-Resolution Nanoscopy Studies of HIV-1 Env Distribution and Mobility. Cell Chem Biol 24 (5):635-645 e635. doi:10.1016/j.chembiol.2017.04.007

22. Peng K, Muranyi W, Glass B, Laketa V, Yant SR, Tsai L, Cihlar T, Muller B, Krausslich HG (2014) Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. eLife 3. doi:10.7554/eLife.04114

23. Sbalzarini IF, Koumoutsakos P (2005) Feature point tracking and trajectory analysis for video imaging in cell biology. J Struct Biol 151 (2):182-195

24. Helmuth JA, Burckhardt CJ, Koumoutsakos P, Greber UF, Sbalzarini IF (2007) A novel supervised trajectory segmentation algorithm identifies distinct types of human adenovirus motion in host cells. J Struct Biol 159 (3):347-358. doi:10.1016/j.jsb.2007.04.003

25. Engelke MF, Burckhardt CJ, Morf MK, Greber UF (2011) The dynactin complex enhances the speed of microtubule-dependent motions of adenovirus both towards and away from the nucleus. Viruses 3 (3):233-253

26. Burckhardt CJ, Suomalainen M, Schoenenberger P, Boucke K, Hemmi S, Greber UF (2011) Drifting motions of the adenovirus receptor CAR and immobile integrins initiate virus uncoating and membrane lytic protein exposure. Cell Host Microbe 10 (2):105-117. doi:10.1016/j.chom.2011.07.006

27. Ewers H, Smith AE, Sbalzarini IF, Lilie H, Koumoutsakos P, Helenius A (2005) Single-particle tracking of murine polyoma virus-like particles on live cells and artificial membranes. Proc Natl Acad Sci U S A 102 (42):15110-15115

28. Yamauchi Y, Boukari H, Banerjee I, Sbalzarini IF, Horvath P, Helenius A (2011) Histone deacetylase 8 is required for centrosome cohesion and influenza A virus entry. PLoS Pathog 7 (10):e1002316. doi:10.1371/journal.ppat.1002316

29. Helmuth JA, Burckhardt CJ, Greber UF, Sbalzarini IF (2009) Shape reconstruction of subcellular structures from live cell fluorescence microscopy images. J Struct Biol 167 (1):1-10. doi:10.1016/j.jsb.2009.03.017

30. Helmuth JA, Paul G, Sbalzarini IF (2010) Beyond co-localization: inferring spatial interactions between sub-cellular structures from microscopy images. BMC Bioinformatics 11:372. doi:10.1186/1471-2105-11-372

31. Bykov YS, Cortese M, Briggs JA, Bartenschlager R (2016) Correlative light and electron microscopy methods for the study of virus-cell interactions. FEBS Lett 590 (13):1877-1895. doi:10.1002/1873-3468.12153

32. Lagache T, Lang G, Sauvonnet N, Olivo-Marin JC (2013) Analysis of the spatial organization of molecules with robust statistics. PLoS One 8 (12):e80914. doi:10.1371/journal.pone.0080914

33. Shivanandan A, Radenovic A, Sbalzarini IF (2013) MosaicIA: an ImageJ/Fiji plugin for spatial pattern and interaction analysis. BMC Bioinformatics 14:349. doi:10.1186/1471-2105-14-349

34. Giesen C, Wang HA, Schapiro D, Zivanovic N, Jacobs A, Hattendorf B, Schuffler PJ, Grolimund D, Buhmann JM, Brandt S, Varga Z, Wild PJ, Gunther D, Bodenmiller B (2014) Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nat Methods 11 (4):417-422. doi:10.1038/nmeth.2869

35. Zenklusen D, Singer RH (2010) Analyzing mRNA expression using single mRNA resolution fluorescent in situ hybridization. Methods Enzymol 470:641-659. doi:10.1016/S0076-6879(10)70026-4

36. Marsh M, Helenius A (2006) Virus entry: open sesame. Cell 124 (4):729-740

37. Flatt JW, Greber UF (2017) Viral mechanisms for docking and delivering at nuclear pore complexes. Semin Cell Dev Biol 68:59-71. doi:10.1016/j.semcdb.2017.05.008

38. Yamauchi Y, Greber UF (2016) Principles of Virus Uncoating: Cues and the Snooker Ball. Traffic 17 (6):569-592. doi:10.1111/tra.12387

39. Sodhi A, Montaner S, Gutkind JS (2004) Viral hijacking of G-protein-coupled-receptor signalling networks. Nat Rev Mol Cell Biol 5 (12):998-1012. doi:10.1038/nrm1529

40. Stertz S, Shaw ML (2011) Uncovering the global host cell requirements for influenza virus replication via RNAi screening. Microbes and infection / Institut Pasteur 13 (5):516-525. doi:10.1016/j.micinf.2011.01.012

41. Franceschini A, Meier R, Casanova A, Kreibich S, Daga N, Andritschke D, Dilling S, Ramo P, Emmenlauer M, Kaufmann A, Conde-Alvarez R, Low SH, Pelkmans L, Helenius A, Hardt WD, Dehio C, von Mering C (2014) Specific inhibition of diverse pathogens in human cells by synthetic microRNA-like oligonucleotides inferred from RNAi screens. Proc Natl Acad Sci U S A 111 (12):4548-4553. doi:10.1073/pnas.1402353111

42. Snijder B, Sacher R, Ramo P, Liberali P, Mench K, Wolfrum N, Burleigh L, Scott CC, Verheije MH, Mercer J, Moese S, Heger T, Theusner K, Jurgeit A, Lamparter D, Balistreri G, Schelhaas M, De Haan CA, Marjomaki V, Hyypia T, Rottier PJ, Sodeik B, Marsh M, Gruenberg J, Amara A, Greber U, Helenius A, Pelkmans L (2012) Single-cell

analysis of population context advances RNAi screening at multiple levels. Mol Syst Biol 8:579. doi:10.1038/msb.2012.9

43. Mercer J, Snijder B, Sacher R, Burkard C, Bleck CK, Stahlberg H, Pelkmans L, Helenius A (2012) RNAi screening reveals proteasome- and Cullin3-dependent stages in vaccinia virus infection. Cell reports 2 (4):1036-1047. doi:10.1016/j.celrep.2012.09.003

44. Meier R, Franceschini A, Horvath P, Tetard M, Mancini R, von Mering C, Helenius A, Lozach PY (2014) Genome-Wide Small Interfering RNA Screens Reveal VAMP3 as a Novel Host Factor Required for Uukuniemi Virus Late Penetration. Journal of Virology 88 (15):8565-8578. doi:10.1128/Jvi.00388-14

45. Green VA, Pelkmans L (2016) A Systems Survey of Progressive Host-Cell Reorganization during Rotavirus Infection. Cell Host Microbe 20 (1):107-120. doi:10.1016/j.chom.2016.06.005

46. Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, Moulton HM, DeJesus P, Che J, Mulder LC, Yanguez E, Andenmatten D, Pache L, Manicassamy B, Albrecht RA, Gonzalez MG, Nguyen Q, Brass A, Elledge S, White M, Shapira S, Hacohen N, Karlas A, Meyer TF, Shales M, Gatorano A, Johnson JR, Jang G, Johnson T, Verschueren E, Sanders D, Krogan N, Shaw M, Konig R, Stertz S, Garcia-Sastre A, Chanda SK (2015) Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding. Cell Host Microbe 18 (6):723-735. doi:10.1016/j.chom.2015.11.002

47. Ramo P, Drewek A, Arrieumerlou C, Beerenwinkel N, Ben-Tekaya H, Cardel B, Casanova A, Conde-Alvarez R, Cossart P, Csucs G, Eicher S, Emmenlauer M, Greber U, Hardt WD, Helenius A, Kasper C, Kaufmann A, Kreibich S, Kuhbacher A, Kunszt P, Low SH, Mercer J, Mudrak D, Muntwiler S, Pelkmans L, Pizarro-Cerda J, Podvinec M, Pujadas E, Rinn B, Rouilly V, Schmich F, Siebourg-Polster J, Snijder B, Stebler M, Studer G, Szczurek E, Truttmann M, von Mering C, Vonderheit A, Yakimovich A, Buhlmann P, Dehio C (2014) Simultaneous analysis of large-scale RNAi screens for pathogen entry. BMC Genomics 15 (1):1162. doi:10.1186/1471-2164-15-1162

48. Staring J, von Castelmur E, Blomen VA, van den Hengel LG, Brockmann M, Baggen J, Thibaut HJ, Nieuwenhuis J, Janssen H, van Kuppeveld FJ, Perrakis A, Carette JE, Brummelkamp TR (2017) PLA2G16 represents a switch between entry and clearance of Picornaviridae. Nature 541 (7637):412-416. doi:10.1038/nature21032

49. Elling U, Wimmer RA, Leibbrandt A, Burkard T, Michlits G, Leopoldi A, Micheler T, Abdeen D, Zhuk S, Aspalter IM, Handl C, Liebergesell J, Hubmann M, Husa AM, Kinzer M, Schuller N, Wetzel E, van de Loo N, Martinez JAZ, Estoppey D, Riedl R, Yang F, Fu B, Dechat T, Ivics Z, Agu CA, Bell O, Blaas D, Gerhardt H, Hoepfner D, Stark A, Penninger JM (2017) A reversible haploid mouse embryonic stem cell biobank resource for functional genomics. Nature. doi:10.1038/nature24027

50. Siebourg-Polster J, Mudrak D, Emmenlauer M, Ramo P, Dehio C, Greber U, Frohlich H, Beerenwinkel N (2015) NEMix: Single-cell Nested Effects Models for Probabilistic Pathway Stimulation. PLoS Comput Biol 11 (4):e1004078. doi:10.1371/journal.pcbi.1004078

51. Sbalzarini IF (2013) Modeling and simulation of biological systems from image data. Bioessays 35 (5):482-490. doi:10.1002/bies.201200051

52. Zhao G, Perilla JR, Yufenyuy EL, Meng X, Chen B, Ning J, Ahn J, Gronenborn AM, Schulten K, Aiken C, Zhang P (2013) Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. Nature 497 (7451):643-646. doi:10.1038/nature12162

53. Schur FK, Obr M, Hagen WJ, Wan W, Jakobi AJ, Kirkpatrick JM, Sachse C, Krausslich HG, Briggs JA (2016) An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. Science 353 (6298):506-508. doi:10.1126/science.aaf9620

54. Larsson DS, Liljas L, van der Spoel D (2012) Virus capsid dissolution studied by microsecond molecular dynamics simulations. PLoS Comput Biol 8 (5):e1002502. doi:10.1371/journal.pcbi.1002502

55. Rapaport DC (2004) Self-assembly of polyhedral shells: a molecular dynamics study. Phys Rev E Stat Nonlin Soft Matter Phys 70 (5 Pt 1):051905. doi:10.1103/PhysRevE.70.051905

56. Freddolino PL, Arkhipov AS, Larson SB, McPherson A, Schulten K (2006) Molecular dynamics simulations of the complete satellite tobacco mosaic virus. Structure 14 (3):437-449. doi:10.1016/j.str.2005.11.014

57. Krishna V, Ayton GS, Voth GA (2010) Role of protein interactions in defining HIV-1 viral capsid shape and stability: a coarse-grained analysis. Biophys J 98 (1):18-26. doi:10.1016/j.bpj.2009.09.049

58. Arkhipov A, Freddolino PL, Schulten K (2006) Stability and dynamics of virus capsids described by coarse-grained modeling. Structure 14 (12):1767-1777. doi:10.1016/j.str.2006.10.003

59. Kim MK, Jernigan RL, Chirikjian GS (2003) An elastic network model of HK97 capsid maturation. J Struct Biol 143 (2):107-117

60. Zheng W, Liao JC, Brooks BR, Doniach S (2007) Toward the mechanism of dynamical couplings and translocation in hepatitis C virus NS3 helicase using elastic network model. Proteins 67 (4):886-896. doi:10.1002/prot.21326

61. English TJ, Hammer DA (2005) The effect of cellular receptor diffusion on receptor-mediated viral binding using Brownian adhesive dynamics (BRAD) simulations. Biophys J 88 (3):1666-1675. doi:10.1529/biophysj.104.047043

62. Szklarczyk OM, Gonzalez-Segredo N, Kukura P, Oppenheim A, Choquet D, Sandoghdar V, Helenius A, Sbalzarini IF, Ewers H (2013) Receptor concentration and diffusivity control multivalent binding of Sv40 to membrane bilayers. PLoS Comput Biol 9 (11):e1003310. doi:10.1371/journal.pcbi.1003310

63. Yakimovich A, Gumpert H, Burckhardt CJ, Lutschg VA, Jurgeit A, Sbalzarini IF, Greber UF (2012) Cell-free transmission of human adenovirus by passive mass transfer in cell culture simulated in a computer model. J Virol 86 (18):10123–10137. doi:10.1128/JVI.01102-12

64. Dinh AT, Theofanous T, Mitragotri S (2005) A model for intracellular trafficking of adenoviral vectors. Biophys J 89 (3):1574-1588. doi:10.1529/biophysj.105.059477

65. Yakimovich A, Yakimovich Y, Schmid M, Mercer J, Sbalzarini IF, Greber UF (2016) Infectio: a Generic Framework for Computational Simulation of Virus Transmission between Cells. mSphere 1 (1). doi:10.1128/mSphere.00078-15

66. Chojnacki J, Staudt T, Glass B, Bingen P, Engelhardt J, Anders M, Schneider J, Muller B, Hell SW, Krausslich HG (2012) Maturation-dependent HIV-1 surface protein redistribution revealed by fluorescence nanoscopy. Science 338 (6106):524-528. doi:10.1126/science.1226359

67. Sadiq SK (2016) Reaction-diffusion basis of retroviral infectivity. Philos Trans A Math Phys Eng Sci 374 (2080). doi:10.1098/rsta.2016.0148

68. Frank GA, Narayan K, Bess JW, Jr., Del Prete GQ, Wu X, Moran A, Hartnell LM, Earl LA, Lifson JD, Subramaniam S (2015) Maturation of the HIV-1 core by a non-diffusional phase transition. Nat Commun 6:5854. doi:10.1038/ncomms6854

69. Mattei S, Glass B, Hagen WJ, Krausslich HG, Briggs JA (2016) The structure and flexibility of conical HIV-1 capsids determined within intact virions. Science 354 (6318):1434-1437. doi:10.1126/science.aah4972

70. Srivastava R, You L, Summers J, Yin J (2002) Stochastic vs. deterministic modeling of intracellular viral kinetics. J Theor Biol 218 (3):309-321

71. Konnyu B, Sadiq SK, Turanyi T, Hirmondo R, Muller B, Krausslich HG, Coveney PV, Muller V (2013) Gag-Pol processing during HIV-1 virion maturation: a systems biology approach. PLoS Comput Biol 9 (6):e1003103. doi:10.1371/journal.pcbi.1003103

72. Gazzola M, Burckhardt CJ, Bayati B, Engelke M, Greber UF, Koumoutsakos P (2009) A stochastic model for microtubule motors describes the in vivo cytoplasmic transport of human adenovirus. PLoS Comput Biol 5 (12):e1000623. doi:10.1371/journal.pcbi.1000623

73. Bremner KH, Scherer J, Yi J, Vershinin M, Gross SP, Vallee RB (2009) Adenovirus transport via direct interaction of cytoplasmic dynein with the viral capsid hexon subunit. Cell Host Microbe 6 (6):523-535. doi:10.1016/j.chom.2009.11.006

74. Labouesse C, Gabella C, Meister JJ, Vianay B, Verkhovsky AB (2016) Microsurgery-aided in-situ force probing reveals extensibility and viscoelastic properties of individual stress fibers. Scientific reports 6:23722. doi:10.1038/srep23722

75. Martin-Fernandez M, Longshaw SV, Kirby I, Santis G, Tobin MJ, Clarke DT, Jones GR (2004) Adenovirus Type-5 Entry and Disassembly Followed in Living Cells by FRET, Fluorescence Anisotropy, and FLIM. Biophys J 87 (2):1316-1327

76. Greber UF, Willetts M, Webster P, Helenius A (1993) Stepwise dismantling of adenovirus 2 during entry into cells. Cell 75 (3):477-486. doi:0092-8674(93)90382-Z [pii]

77. Nakano MY, Boucke K, Suomalainen M, Stidwill RP, Greber UF (2000) The first step of adenovirus type 2 disassembly occurs at the cell surface, independently of endocytosis and escape to the cytosol. J Virol 74 (15):7085-7095

78. Greber UF (2016) Virus and Host Mechanics Support Membrane Penetration and Cell Entry. J Virol 90 (8):3802-3805. doi:10.1128/JVI.02568-15

79. Snijder J, Reddy VS, May ER, Roos WH, Nemerow GR, Wuite GJ (2013) Integrin and defensin modulate the mechanical properties of adenovirus. J Virol 87 (5):2756-2766. doi:10.1128/JVI.02516-12

80. Ortega-Esteban A, Condezo GN, Perez-Berna AJ, Chillon M, Flint SJ, Reguera D, San Martin C, de Pablo PJ (2015) Mechanics of Viral Chromatin Reveals the Pressurization of Human Adenovirus. ACS Nano 9 (11):10826-10833. doi:10.1021/acsnano.5b03417

81. Ortega-Esteban A, Bodensiek K, San Martin C, Suomalainen M, Greber UF, de Pablo PJ, Schaap IA (2015) Fluorescence Tracking of Genome Release during Mechanical Unpacking of Single Viruses. ACS Nano 9 (11):10571-10579. doi:10.1021/acsnano.5b03020

82. Ortega-Esteban A, Perez-Berna AJ, Menendez-Conejero R, Flint SJ, San Martin C, de Pablo PJ (2013) Monitoring dynamics of human adenovirus disassembly induced by mechanical fatigue. Scientific reports 3:1434. doi:10.1038/srep01434

83. Kusumi A, Nakada C, Ritchie K, Murase K, Suzuki K, Murakoshi H, Kasai RS, Kondo J, Fujiwara T (2005) Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. Annu Rev Biophys Biomol Struct 34:351-378

84. Iizuka N, Oka M, Yamada-Okabe H, Mori N, Tamesa T, Okada T, Takemoto N, Tangoku A, Hamada K, Nakayama H, Miyamoto T, Uchimura S, Hamamoto Y (2002) Comparison of gene expression profiles between hepatitis B virus- and hepatitis C virus-infected hepatocellular carcinoma by oligonucleotide microarray data on the basis of a supervised learning method. Cancer Res 62 (14):3939-3944

85. Snijder B, Sacher R, Ramo P, Damm EM, Liberali P, Pelkmans L (2009) Population context determines cell-to-cell variability in endocytosis and virus infection. Nature 461 (7263):520-523. doi:10.1038/nature08282

86. Helmuth JA, Sbalzarini IF (2009) Deconvolving active contours for fluorescence microscopy images. Lecture Notes in Computer Science. In: Proc. Intl. Symp. Visual Computing (ISVC), vol 5875. Las Vegas, USA, pp 544–553

87. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, S. T (2008) Global Sensitivity Analysis: The Primer. . 1st edition edn. Wiley Interscience,

88. Sbalzarini IF, Mezzacasa A, Helenius A, Koumoutsakos P (2005) Effects of organelle shape on fluorescence recovery after photobleaching. Biophys J 89 (3):1482-1492

89. Kjellström G, Taxen L (1981) Stochastic optimization in system design. IEEE Trans Circ and Syst 28:702–715

90. Thach PT (1988) The design centering problem as a D.C. programming problem. Mathematical Programming 41:229–248

91. Asmus J, Muller CL, Sbalzarini IF (2017) L p -Adaptation: Simultaneous Design Centering and Robustness Estimation of Electronic and Biological Systems. Scientific reports 7 (1):6660. doi:10.1038/s41598-017-03556-5

92. Reddy T, Shorthouse D, Parton DL, Jefferys E, Fowler PW, Chavent M, Baaden M, Sansom MS (2015) Nothing to sneeze at: a dynamic and integrative computational model of an influenza A virion. Structure 23 (3):584-597. doi:10.1016/j.str.2014.12.019

93. Miao Y, Fu R, Zhou HX, Cross TA (2015) Dynamic Short Hydrogen Bonds in Histidine Tetrad of Full-Length M2 Proton Channel Reveal Tetrameric Structural Heterogeneity and Functional Mechanism. Structure 23 (12):2300-2308. doi:10.1016/j.str.2015.09.011

94. Takeda M, Pekosz A, Shuck K, Pinto LH, Lamb RA (2002) Influenza a virus M2 ion channel activity is essential for efficient replication in tissue culture. J Virol 76 (3):1391-1399

95. Shimbo K, Brassard DL, Lamb RA, Pinto LH (1996) Ion selectivity and activation of the M2 ion channel of influenza virus. Biophys J 70 (3):1335-1346. doi:10.1016/S0006-3495(96)79690-X

96. Martin K, Helenius A (1991) Nuclear transport of influenza virus ribonucleoproteins: the viral matrix protein (M1) promotes export and inhibits import. Cell 67 (1):117-130

97. Fontana J, Steven AC (2013) At low pH, influenza virus matrix protein M1 undergoes a conformational change prior to dissociating from the membrane. J Virol 87 (10):5621-5628. doi:10.1128/JVI.00276-13

98. Ison MG (2011) Antivirals and resistance: influenza virus. Current opinion in virology 1 (6):563-573. doi:10.1016/j.coviro.2011.09.002

99. Greber UF (2014) How cells tune viral mechanics--insights from biophysical measurements of influenza virus. Biophys J 106 (11):2317-2321. doi:10.1016/j.bpj.2014.04.025

100. Cross TA, Dong H, Sharma M, Busath DD, Zhou HX (2012) M2 protein from influenza A: from multiple structures to biophysical and functional insights. Current opinion in virology 2 (2):128-133. doi:10.1016/j.coviro.2012.01.005

101. Bright RA, Shay DK, Shu B, Cox NJ, Klimov AI (2006) Adamantane resistance among influenza A viruses isolated early during the 2005-2006 influenza season in the United States. JAMA 295 (8):891-894. doi:10.1001/jama.295.8.joc60020

102. Gleed ML, Busath DD (2015) Why bound amantadine fails to inhibit proton conductance according to simulations of the drug-resistant influenza A M2 (S31N). J Phys Chem B 119 (3):1225-1231. doi:10.1021/jp508545d

103. Wright AK, Batsomboon P, Dai J, Hung I, Zhou HX, Dudley GB, Cross TA (2016) Differential Binding of Rimantadine Enantiomers to Influenza A M2 Proton Channel. J Am Chem Soc 138 (5):1506-1509. doi:10.1021/jacs.5b13129

104. Pinto LH, Holsinger LJ, Lamb RA (1992) Influenza virus M2 protein has ion channel activity. Cell 69 (3):517-528

105. Holsinger LJ, Lamb RA (1991) Influenza virus M2 integral membrane protein is a homotetramer stabilized by formation of disulfide bonds. Virology 183 (1):32-34

106. Leiding T, Wang J, Martinsson J, DeGrado WF, Arskold SP (2010) Proton and cation transport activity of the M2 proton channel from influenza A virus. Proc Natl Acad Sci U S A 107 (35):15409-15414. doi:10.1073/pnas.1009997107

107. Stauffer S, Feng Y, Nebioglu F, Heilig R, Picotti P, Helenius A (2014) Stepwise priming by acidic pH and high K+ is required for efficient uncoating of influenza A virus cores after penetration. J Virol. doi:10.1128/JVI.01430-14

108. Chizhmakov IV, Geraghty FM, Ogden DC, Hayhurst A, Antoniou M, Hay AJ (1996) Selective proton permeability and pH regulation of the influenza virus M2 channel expressed in mouse erythroleukaemia cells. J Physiol 494 (Pt 2):329-336

109. Scott CC, Gruenberg J (2011) Ion flux and the function of endosomes and lysosomes: pH is just the start: the flux of ions across endosomal membranes influences endosome function not only through regulation of the luminal pH. BioEssays : news and reviews in molecular, cellular and developmental biology 33 (2):103-110. doi:10.1002/bies.201000108

110. Suomalainen M, Greber UF (2013) Uncoating of non-enveloped viruses. Current opinion in virology 3:27-33. doi:10.1016/j.coviro.2012.12.004

111. Kilcher S, Mercer J (2015) DNA virus uncoating. Virology. doi:10.1016/j.virol.2015.01.024

112. Polyansky AA, Ramaswamy R, Volynsky PE, Sbalzarini IF, Marrink SJ, Efremov RG (2010) Antimicrobial Peptides Induce Growth of Phosphatidylglycerol Domains in a Model Bacterial Membrane. Journal of Physical Chemistry Letters 1 (20):3108-3111. doi:DOI: 10.1021/jz101163e

113. Luisoni S, Suomalainen M, Boucke K, Tanner LB, Wenk MR, Guan XL, Grzybek M, Coskun U, Greber UF (2015) Co-option of Membrane Wounding Enables Virus Penetration into Cells. Cell Host Microbe 18 (1):75-85. doi:10.1016/j.chom.2015.06.006

114. Luisoni S, Greber UF (2016) Biology of Adenovirus Cell Entry – Receptors, Pathways, Mechanisms. In: Curiel D (ed) Adenoviral Vectors for Gene Therapy. 2nd edn. Academic Press, Elsevier, London, UK, pp 27-58